Big Data Recommender System for Encouraging Purchases in New Places Taking into Account Demographics

Hugo Alatrista-Salas[1][0000−0001−5252−4728],
Isaías Hoyos[1], Ana Luna[1][0000−0002−1359−2469], and
Miguel Nunez-del-Prado[1][0000−0001−7997−1739]

[1]Universidad del Pacífico (Lima, Perú)
{h.alatristas,i.hoyosl,ae.lunaa,m.nunezdelpradoc}@up.edu.pe

Abstract. Recommendation systems have gained popularity in recent years. Among them, the best known are those that select products in stores, movies, videos, music, books, among others. The companies, and in particular, the banking entities are the most interested in implementing these types of techniques to maximize the purchases of potential clients. For this, it is necessary to process a large amount of historical data of the users and convert them into useful information that allows predicting the products of interest for the user and the company. In this article, we analyze two essential problems when using systems, one of which is to suggest products of one commerce to those who have never visited that place, and the second is how to prioritize the order in which users buy certain products or services. To confront these drawbacks, we propose a process that combines two models: latent factor and demographic similarity. To test our proposal, we have used a database with approximately 65 million banking transactions. We validate our methodology, achieving an increase in the average consumption in the selected sample.

Keywords: Recommender system · Latent Factor · Consumption Patterns· Demographic Vector.

1 Introduction

Nowadays, there is a large number of products and services that the markets offer. In consequence, many users have difficulties in choosing the items they need in an easy and fast way. This brings a problem for companies that make profits from customers' purchases. Thus, business prioritizes satisfaction to preserve loyalty and close relationship with the clients. Hence, companies face new challenges to elaborate strategies for keeping the client's loyalty. These new strategies need to take into account the buyer's data, commercial activity, preferences, working place, among others. In this context, companies collect large amounts of data from customers to use them as a source of information. They transform data into useful insights that allow customers acquiring products based on their preferences. Therefore recommendation systems help customers to filter information in a personalized and transparent way.

One of the best-known examples of these systems is Netflix. The popular platform for movies and series streaming analyzes large volumes of data, such as what a user watches, at what time, for how long and what device is using. After analyzing the data, the Netflix algorithms recommend the user what to see based on the analyzed data, presenting the recommendations on the home page [11]. The objective of the recommendation algorithms that Netflix uses is to compensate for the low decision power of the human being before a wide range of options [16].

Gallego and Huecas [9] describe another example of a recommendation system in an economic context implemented in Spain. The objective is to recommend commerces taking into account the spatial and temporal dynamics of clients. This system seeks to recommend entities where bank customers can pay with their bank cards. To accomplish this, the bank uses large volumes of customers' data, like the transactions made by clients and the shops they visited. The recommendation looks forward to using the client's credit card and knowing where to use it according to the existing spatial information. But, the main difference in our research approach compared to others is that we add the similarity method that takes into account the demographic characteristics of the clients to prioritize the most similar shops. In this context, the present article describes the process of building an innovative recommendation system. The objective of our proposal is to encourage the use of credit and debit cards from a financial institution having three million users of credit and debit cards in July of 2017.

As a direct consequence of promoting the use of credit and debit cards, the average of the customer's purchase tickets increased. This particular point is one of the motivations of this work. Our aim was to generate a recommendation system that enhances the use of cards in commerces that the clients have never visited before.

The rest of the article is structured as follows. Section 2 describes the state of the art, Section 3 introduces the theoretical foundations of our proposal. Then, Section 4 describes the experiments and presents the obtained results, while Section 5 shows the validation of the results. Finally, Section 6 outlines the conclusions and future works.

2 Related Works

There are several studies on recommendation systems, where four approaches are identified and differentiated: *content-based*, *collaborative filtering*, *demographic filtering*, and *hybrid*.

The first one considers the profiles of the users and the decisions that they made in the past to recommend a particular item. For this, the similarity between articles is used to suggest those closest to the user's preference [4], [17]. The clients' preference description is used to distinguish one article from another [2]. Then, attributes are compared with the user's profile and items with the highest degree of similarity are recommended. Content-based recommendation engine needs existing profiles of users that has information about a user and their taste. So, it works on the basis of item attributes. However, gathering descriptions requires external information that may not be available or that is difficult to collect. Besides, the recommendations tend to overspecialize the products or items. Thus, the obtained recommendations are not novel nor useful. An example of the use of the content-based approach is found in the Genome Musical Project, in which songs based on 400 musical characteristics that capture the musical "identity" of a song are recommended [8]. In this approach, heuristic recommendations were generated using a traditional information retrieval technique known as cosine similarity.

The second approach is *collaborative filtering*, which takes into account the past behavior of the user, namely items valuations or previous acquisitions. This treatment uses the known preferences of the client to predict the unknown ones to make a recommendation. It works finding users in a community that shares appreciations, so they will have similar tastes. Such users build a group or a so-called neighborhood. A user gets recommendations for those items that the user has not rated before but was positively rated by users in his/her neighborhood. So, the collaborative filtering also uses user behavior in addition to product attributes and unlike the content-based recommendation engine, you can recommend articles in other categories that the user never viewed before. There exist two methods in this approach: 1) the *neighborhood method*, based on the similarity between the items and users [1]; and 2) the model of *latent factor*, which consists of representing the users and the items through factors or characteristics [19]. For instance, Koren *et al.* [13], developed a recommendation system for the Netflix competition based on latent factor models, matrix factorization and the use of alternating least squares (ALS). About a couple of years ago, Covington*et al.* [8] described the Youtube recommendation system using deep neural networks and neighborhood methods to process searches. In banking, recommendation systems have been developed with information on customers' use of debit or credit cards, taking into account the clients' spatial and temporal contexts as well as their similarity with other clients when visiting commerces [9].

The third recommendation approach is *demographic filtering*. The main idea behind this approach is that users with similar characteristics and attributes will have similar preferences [4]. Its function is to categorize the user according to his attributes and create recommendations based on demographic variables (such as age and sex, among others) [15]. For example, Al-Shamri [3] proposes five different metrics to measure demographic similarity for recommendation systems. (1) *Mixed Profiling Approach* considers the similarity of each attribute independently and then it aggregates the similarity of each one in a single score. (2) *Categorical Profiling Approach* unifies those attributes that yield a similarity value equal to 1 when their categories are equivalent. (3) *Fuzzy Profiling Approach* exploits the vague nature of the age attribute. Thus, two users with close ages are 100% similar. (4) *Cascaded Profiling Approach* considers age as the only factor for obtaining user similarity in the beginning. Then, the system elects a big set of neighbors relying on age similarity. This more significant set is the input to another system, which takes into account both age and gender attributes. The output of this system is a new smaller set of neighbors. Finally, the system uses age, gender, and occupation attributes. (5) *Single-Attribute Profiling Approach* takes each attribute as a profile. Hence, there is an independent recommender system for each attribute like age, gender, and occupation. To perform experiments, the author used the 100K MovieLens dataset with the Mean Absolute Error, Root Mean Squared Error, Coverage and percentage of the correct predictions metrics for recommendations systems evaluations. He found that the way of outlining users play an essential role in enhancing system performance.

Finally, the literature shows the *hybrid* approach, in which the characteristics of the collaborative filter and the content-based filter are combined, even some authors use recommendation systems based on demographic analysis. Zhao et al. [20] employed demographic information on social media platforms to boost sales. For this purpose, they extracted product and user demographics from online product reviews and social networks built from microblogs for product recommendation.

The authors of the reference [7] investigated and studied the problem that arises in the recommendation systems when the data of the new users are not available. The alternative they propose to minimize this lack of precedents is basing on the user's demographic information.

Additionally, the *hybrid* approaches address the problem of the cold start of the collaborative filter and the over-specialization of content-based techniques. Burke [6] details different methods for linking the existing content-based and collaborative filter approaches by weighting, exchange, mixing, among others. The system needs data to understand the preferences of the users and makes recommendations that express the client's preference towards an item through an explicit or implicit valuation. On the one hand, the explicit valuation indicates the user's choice using a scale, such as assigning 1 to 5 stars to commerce. On the other hand, the implicit valuation is calculated by analyzing the user's behavior in the domain of the system; for example, the amount of time devoted to observing a tab and the amount spent in commerces.

Due to the pros and cons of the methods described above, we propose to combine the factor models and the demographic vector, which allows reordering the recommended shops through the cosine similarity model which is based on linear algebra rather than statistical approach. That model is commonly used to calculate the scores that express how similar users or items are to each other. These scores can then be used as the foundation of user- or item-based recommendation generation.

In this article, the amount of data with which we worked is fixed. However, if the data grows at very high velocity, it is prudent to use complementary methods that can handle such data accurately as well as efficiently in the context of a recommender system. In those cases, Kumar [14] proposes a latent factor model that caters to both accuracy and efficiency by reducing the complexity of the model without compromising on accuracy.

3 Recommendation System Process

Our proposal is divided into six phases, as shown in Figure 1. First, we obtain a database that contains clients' credit and debit card transactions. Then, online transactions that do not belong to commerce are filtered. After that, remaining transactions are divided according to the area (district) where the commerce belongs. Both data and the filtering process will be described in section 4. Then, using latent factor, alternating matrices factorization, and least squares, the recommender system estimates the unknown valuations of the clients. These techniques are detailed in subsections 3.1, 3.2 and 3.3, respectively. Therefore, all the *ratings* of the clients are obtained per commerce. Once ratings are collected, we use demographic characteristics to refine recommended commerces order. Finally, we plotted them in cartography.

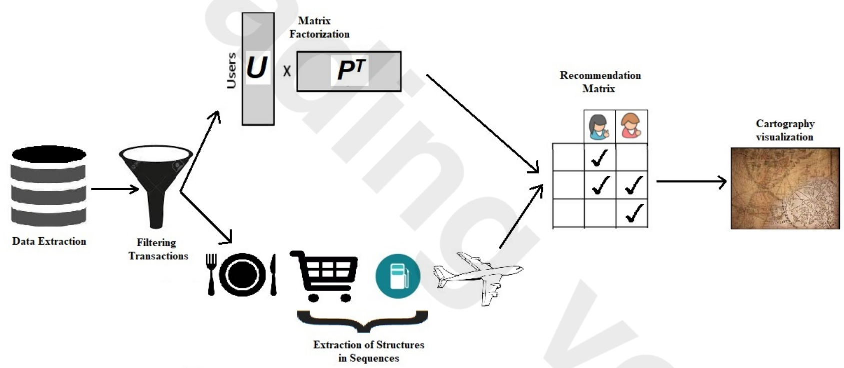


Fig. 1. Operation recommender system scheme.

3.1 Latent factor models

The latent factor models characterize the items and the users by factors inferred from the *ratings* to explain items valuations [19]. Matrix factorization techniques are critical to these models, as they find the values of the latent factor to estimate unknown valuations. Each user and each item is modeled with an individual vector of specific features. Using the scalar product of these individual vectors, we would estimate the valuation of a user to an item. In the case of an item, the feature vector indicates its factor as a result of an assigned score. The higher the score is, the better is the item characterized. For example, a movie that has a high rating on the "Drama" feature will be a dramatic genre movie.

3.2 Matrix factorization

Based on the latent factor models, the matrix factorization represents the interactions between users and items as a scalar product to obtain the unknown ratings. This approach is the most accurate when dealing with the sparse data in the valuations matrices [5]. Given $u_i \in U$ and $m_j \in M$, which represent the feature vector of the i-th user and the j-th item, respectively, the internal product $u_i^T m_j$ is the interaction between the respective users and items, where u_i^T is the transpose of u_i. Due to the item-user interaction is given by the user's valuation, the Equation 1 expresses the i-th user's valuation to the j-th item.

$$r_{ij} = u_i^T m_j \tag{1}$$

To calculate the values for the matrices U and M, the matrix R must be approximated by minimizing the loss function. Thus, given a pair of matrices U and M, the total loss of the model will be the sum of all the losses in all known valuations, which results in the mean square error (MSE, *c.f.*, Equation 2).

$$f(R, U, M) = \frac{1}{n} \sum_{i,j} (r_{i,j} - u_i^T \times m_j)^2. \tag{2}$$

When valuing the error, one should avoid counting items with no valuations. For this reason, in the Equation 2, we added a matrix W of weights with the same dimension as the matrix R. If $r_{i,j}$ exists, then $w_{i,j}$ takes the value of 1, and 0 otherwise. Therefore, the loss function can be rewritten as shown in Equation 3.

$$f(R, W, U, M) = \frac{1}{n} \sum_{i,j} w_{i,j} (r_{i,j} - u_i^T \times m_j)^2. \tag{3}$$

Then, one can formulate the problem of finding the values of U and M that best approximate the matrix R through Equation 4.

$$(U, M) = \underset{(U,M)}{\arg\min}\, f(R, W, U, M), \tag{4}$$

3.3 Alternating Least Squares (ALS)

ALS is an optimization method to solve Equation 3. Since we have a non-convex function, we ignore both the values of U and M. Thus; we must solve Equation 3 for each unknown variable separately [21]. Besides, we have $(i+j) \times k$ unknown parameters that must be calculated, and an R sparse matrix (since not all users score all trades), and when solving the Equation 3 can generate data overfitting. Therefore, we add a term in Equation 3 using the Tikhonov regularization method [10], as observed in Equation 5.

$$f(R,W,U,M) = \sum_{(i,j)} w_{(i,j)}(r_{(i,j)} - u_i^T \times m_j)^2 + \lambda(\sum_i n_{u_i}||u_i||^2 + \sum_j n_{m_j}||m_j||^2), \quad (5)$$

Being n_{u_i} and n_{m_j} the number of valuations that the user i has on the item m_j, respectively. Using Equation 5 as the function to be minimized, the way to proceed using alternating least squares is as follows:

– Initialize M randomly with values between 1 and 0
– Fix M, and derive partially from U
– Fix U, and derive partially from M
– Repeat the 2nd and 3rd step until reaching the pre-established detention criterion.

The detention criterion is set as the square root of the difference between MSE in steps n and $n-1$. In our case, we used 10^{-4} as the threshold for the stop criterion, as suggested in Zhou *et al.* [21].

3.4 Demographic Rearrangement

The first phase of this stage consisted of generating a commerce demographic vector (d_c) composed of the percentage of males and females who purchased in that specific commerce and the percent of the clients in age ranges from 18-24, 25-30, 31-40, 41-50, 51-60, and 61+. Analogously, we characterized the clients (d_u). It is worth noting that the age ranges were found empirically. Once their demographic vectors describe clients and commerce, it is possible to rearrange the order of the recommendation based on the demographic similarity of the commerce d_c to the client d_u. Therefore, to sort these stores, we use Cosine Similarity, detailed in Equation 6. This process allows us to measure the demographic correspondence of a user and the recommended shops. Thus, the closer to one is the cosine, the higher the similarity is.

$$cosim(\overrightarrow{d_c}, \overrightarrow{d_u}) = \frac{(\overrightarrow{d_c} \cdot \overrightarrow{d_u})}{\|\overrightarrow{d_c}\| \times \|\overrightarrow{d_u}\|}, \quad (6)$$

4 Experiments and Results

The data used in this work are associated with banking transactions registered for one year, specifically between June 2016 to July 2017. In total, we recorded 65 085 138 transactions; approximately 80% of them correspond to transactions made with a debit card, while the remaining 20% were made with credit cards. Besides, 25 variables provide information about the customer, the card owner, the commerce of the transaction, among others. For our experiments, the variables *CodClient*, *CodCommerce*, *Category*, and *Amount_Rating* were selected. We calculated the latter using the Equation 7, which is the ratio between the purchase value in certain commerce and the total amount spent by a customer. This variable represents the valuations of the clients towards the commerces.

$$amount\_rating_{i,j} = \frac{amount\_commerce_j}{amount\_total_i} \tag{7}$$

First of all, we grouped the data by regions (districts) in which the commercial transactions were carried out. For each region, the valuation sparse matrix R is generated. As there are many unknown valuations, the matrix R^* is estimated by matrix factorization, and alternating least squares, finding the values of the user matrix U and the items matrix M. In R^*, the valuation of the client i to the commerce j is represented by r_{ij}.

We used the Python programming language and an instance of Elastic Compute Cloud (EC2) from Amazon Web Services (AWS) to process the data. Once we calculated the matrices, we normalized the rows, and then, the n commerces with the highest valuations are recommended.

The bank classified the commerces to find the user's consumption patterns. To do this, it groups related elements and groups them into different categories. Finally, 22 types of classifications were obtained, each of them coded by a symbol, as shown in Table 1.

Table 1. Categories' Table

NAME_CATEGORÍE	**COD**	**NAME_CATEGORÍE**	**COD**
PRODMARKET	a	RESTBAR	b
HEALTH	c	VEHIDER	d
PRIVSHOP	e	ENTRENT	f
FASHCLOT	g	BEAUTY	h
TELECOMM	i	RENTGOODS	j
FINANC	k	ELECTPROD	l
ARTCULT	m	EDUCATION	n
LOCALPROD	o	TRANSLI	p
DIVPROP	q	CLUBMA	r
HOGOFIC	s	PROFDIV	t
INFORM	u	RETMAN	v

To accomplish our work, we first applied the ALS optimization method in the district of Barranca, located in the province of the same name, in the department of Lima. We showed in Figure 2 that after 5 iterations, the MSE obtained by minimizing Equation 5 is 0.0000026. Thus, the R matrix converges to the matrix R^* quickly.

Then, the predicted valuations are in the matrix R^*. Since there are 167 883 clients, 19 949 commerces and more than 3 billion of valuations, Table 2 shows,

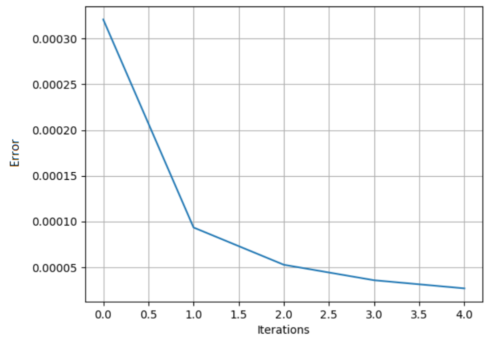


Fig. 2. MSE analyzed in Barranca district

for example, valuations for three clients to three commerces. From this matrix, the *ratings* are expressed using demographic similarity.

Table 2. Predicted valuations Matrix: district of Barranca

	100070254	**100073401**	**101081734**
Bob	0.7935	0.5134	0.2376
Alice	0.8317	0.9205	0.5462
Eve	0.6911	0.4123	0.3433

In the next section, we describe the method to validate the results of the recommender system.

5 Recommendation System Validation

For the recommender verification, a preliminary test and validation were performed on samples corresponding to some districts of Lima city.

5.1 Preliminary test

For the recommendation system test, a survey on the acceptance of the results was conducted with 10 credit and debit cards clients. We computed the recommended commerces by estimating the valuations of 149 362 clients. We divided these commerces into two types: 1) recurring, places that the client had previously visited, and 2) recommended new places for the client. The surveys measured the acceptance of the recommendation by customers for different categories and commerces. In summary, the study was personalized for each user based on their historical consumption and recommendation of new commerces.

Table 3 presents the percentage of recommendations of commerces and categories, both recurrent and novel, in which the client would accept to make some consumption.

Table 3. Test results

	Recurrent	**New**
category	77.0%	76.0%
commerce	84.5%	70.0%

In the case of the commercial category, the acceptance is 77% and 76% for recurring and new commerces, respectively. In the case of merchants *per se*, the acceptance is 84.5% for recurrent commerces and 70% for new commerces (*c.f.*, Table 3). Although the size of the taken sample is small, this allowed us to test the congruence of the recommender in a pilot test. Based on the results, it is possible to see that the recommendation engine is 76% relevant in terms of the new items it suggests, and 70% concerning the new commerces.

This first test was carried out to verify the coherence of the recommendations ratified by the users surveyed. However, for a statistically rigorous validation, the sample size and other factors presented below must be taken into account.

5.2 Validation

For the validation process, we chose four districts of Lima: San Isidro, San Borja, Miraflores, and Independencia, and for each of them, we took a random sample of clients, n = 4500. Due to confidentiality reasons, we can not rebel the exact number of clients by district. However, to give an idea of the order of magnitude of clients, in the Table 4 we see that San Isidro, San Borja, Miraflores, and Independencia represent three, four, five and two times approximately the size of the sample.

District	San Isidro	San Borja	Miraflores	Independencia
Sample size	3n	4n	5n	2n
Sample mean	0.0043	0.0071	0.0080	0.004
Sample population	0.0042	0.0072	0.0079	0.003
Test-t	-0.55	0.25	-0.36	1.72
KLD	0.0009	0.001	0.0018	0.0009
Minimum sample size	410	422	415	400

Table 4. Summary of the sample validation results

To verify that the sample is representative of the population, we carried out three different tests. First, we prove that the samples have the same mean as their respective populations using the test *One Sample t Test* [18]. To perform this test, the confidence of 5% (*i.e.*, α=0.05) was taken with 4499 degrees of freedom and a T value of 1.9605 for a two-tailed test. We raised the null hypothesis which was that the samples mean of the consumption and the population are equal $H_0 : \mu_0 - \mu = 0$. The alternative hypothesis is that these means are different $H_{alt} : \mu_0 - \mu \neq 0$. The calculated values of the *test t* are found in Table 4 and were calculated using Equation 8.

$$t = \frac{\overline{x} - \mu}{\frac{s}{\sqrt{n}}} \tag{8}$$

Where $\overline{x}$ is the population mean, μ is the sample mean, s is the standard deviation, and n is the size of the sample. As a result of the test, the null hypothesis is accepted, so the mean of the sample and population consumption are equal.

The second method uses the divergence of *Kullback-Leibler* to show that from the distribution of the sample p, we can reconstruct the distribution of the population q [12]. For this, we use Equation 9. We compared the distributions in Figure 3.

$$KLD = \sum_{k=1}^{n} p_k \times ln(\frac{(p_k)^2}{q_k}) \tag{9}$$

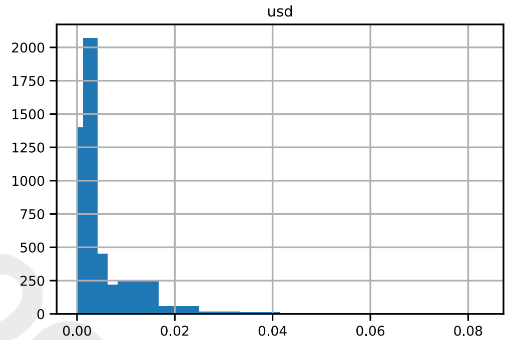
Thus, Table 4 shows the values obtained for the divergence. Note that the values close to zero mean that there is no divergence, that is, the distributions are equal.

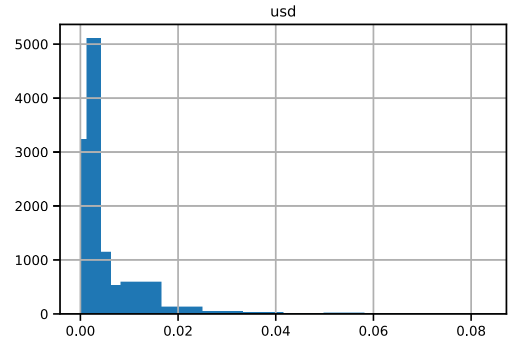
Finally, the size of the sample ($n*$) necessary to have a statistically significant representation was calculated. Knowing the number of clients in each of the districts, we use the formula of Equation 10 to calculate the value of $n*$.

$$n* = \frac{N \times (Z)^2 \times p \times q}{(e)^2 \times (N-1) + (Z)^2 \times p \times q} \tag{10}$$

Where N is the size of the population, Z is the deviation from the average value that we accept to achieve the desired level of confidence (95%), in our case $Z = 1.96$, p is the probability of success or proportion that we expect to find (50%), q represents the probability of failure and e is the maximum margin of error that we admit (5%). To preserve the confidentiality of the exact number of clients, Table 4 was obtained, overestimating the sample size values $n*$. In all cases, the results obtained are an order of magnitude lower than the real value taken for each district.

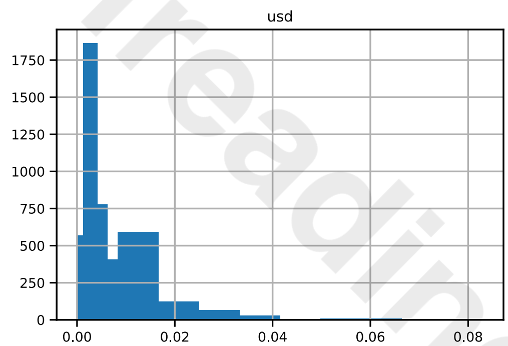
Once it has been demonstrated that the samples taken are statistically representative of their respective populations, we described the results of the recommendation engine. The validation of the engine was carried out using the customers selected as sample, to whom personalized offers were sent in two retail items and restaurants during the first three weeks of March 2018. At the end of that month, the average tickets were compared to the sample and the

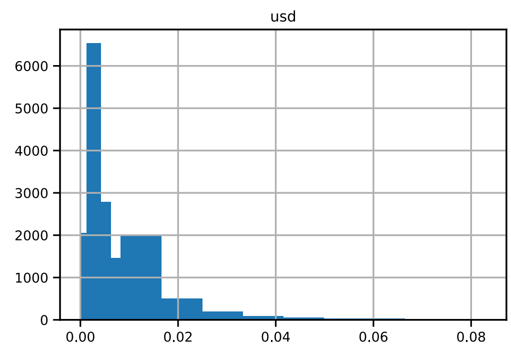




a) San Isidro sample distribution

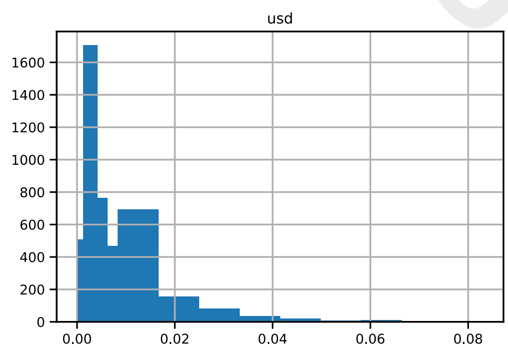
b) San Isidro population distribution

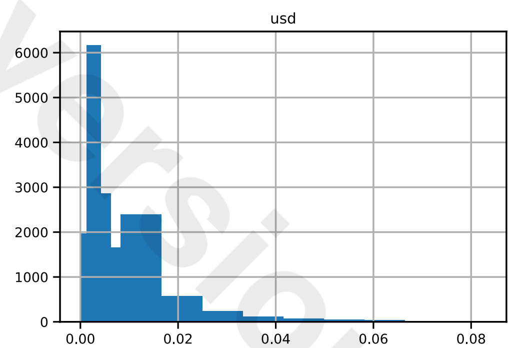




c) San Borja sample distribution

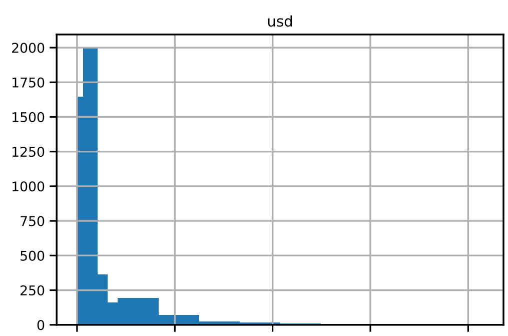
d) San Borja population distribution

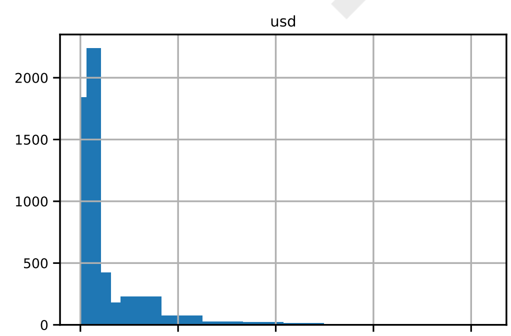




e) Miraflores sample distribution

f) Miraflores population distribution





g) Independencia sample distribution

h)Independencia population distribution

Fig. 3. Samples and populations distributions from different districts.

population without the users in the sample, and it was observed that there was an increase in the average consumer ticket in retailers and restaurants of 10% and 22%, respectively. That meant an increase in expenses from 33.6 USD to 37 USD in retailers and from 27.88 USD to 34.24 USD in restaurants. These results show us the relevance of the recommendation engine and its potential impact on the commerces.

6 Conclusions and Future Works

In this work, we developed a recommendation system for a financial institution. We noted that the amount spent on trade could be used as an implicit valuation, and reflected the preference of the client for different commerces, so the higher the amount paid by the client, the greater is the probability that the client chooses that kind of stores or restaurants.

We took into account the geographic context of the clients when we generated the grouping by area (district), identifying where they made the transactions. The district's geographic division maintained the recommendations within a coherent geographic framework, avoiding recommending commerces in remote areas or places that were visited only once, for example, during the vacation period. Besides, using the measure of cosine similarity, an order of priority was obtained in the recommended commerces. The recommendations are innovative, although it is not ruled out that previously visited commerces are included in the list of recommended commerces.

In relation to future work, we will explore the study and a more detailed density analysis with all the districts of Lima, in order to improve the visualization. In addition, given the high volume of data that we worked on, we are planning to use the integrated Apache Spark *framework* to implement advanced analysis, such as processing data in parallel and working in a shared environment. This can be executed in *clusters* of Elastic Map Reduce (EMR) using MrJob in Hadoop, PySpark or Scala. Also, the privacy of the users and their data must be taken into consideration, with which the recommendations are executed. Consequently, our idea is to implement an additional layer to the recommendation engine to guarantee a certain level of privacy. Finally, we intend to perform a large-scale pilot with ten thousand users to validate the recommender system and have a better measure of real accuracy and not subjective.

Acknowledgements

Authors are grateful for the funding of the Universidad del Pacífico Research Vice-Rector in the grant of the project PY-ESP-0210013216.

References

1. Abdollahpouri, H., Abdollahpouri, A.: An approach for personalization of banking services in multi-channel environment using memory-based collaborative filtering.

In: The 5th Conference on Information and Knowledge Technology. pp. 208–213 (May 2013). https://doi.org/10.1109/IKT.2013.6620066
2. Aggarwal, C.C., et al.: Recommender systems. Springer (2016)
3. Al-Shamri, M.Y.H.: User profiling approaches for demographic recommender systems. Knowledge-Based Systems **100**, 175–187 (2016)
4. Bobadilla, J., Ortega, F., Hernando, A., Gutiérrez, A.: Recommender systems survey. Knowledge-based systems **46**, 109–132 (2013)
5. Bokde, D., Girase, S., Mukhopadhyay, D.: Matrix factorization model in collaborative filtering algorithms: A survey. Procedia Computer Science **49**(Supplement C), 136 – 146 (2015). https://doi.org/https://doi.org/10.1016/j.procs.2015.04.237, http://www.sciencedirect.com/science/article/pii/S1877050915007462, proceedings of 4th International Conference on Advances in Computing, Communication and Control (ICAC3'15)
6. Burke, R.: Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction **12**(4), 331–370 (Nov 2002). https://doi.org/10.1023/A:1021240730564, http://dx.doi.org/10.1023/A:1021240730564
7. Callvik, J., Liu, A.: Using demographic information to reduce the new user problem in recommender systems (2017)
8. Covington, P., Adams, J., Sargin, E.: Deep neural networks for youtube recommendations. In: Proceedings of the 10th ACM Conference on Recommender Systems. pp. 191–198. RecSys '16, New York, NY, USA (2016). https://doi.org/10.1145/2959100.2959190, http://doi.acm.org/10.1145/2959100.2959190
9. Gallego, D., Huecas, G.: An empirical case of a context-aware mobile recommender system in a banking environment. In: Proceedings of the 2012 Third FTRA International Conference on Mobile, Ubiquitous, and Intelligent Computing. pp. 13–20. MUSIC '12, Washington, DC, USA (2012). https://doi.org/10.1109/MUSIC.2012.11, http://dx.doi.org/10.1109/MUSIC.2012.11
10. Golub, G.H., Hansen, P.C., O'Leary, D.P.: Tikhonov regularization and total least squares. SIAM Journal on Matrix Analysis and Applications **21**(1), 185–194 (1999)
11. Gomez-Uribe, C.A., Hunt, N.: The netflix recommender system: Algorithms, business value, and innovation. ACM Trans. Manage. Inf. Syst. **6**(4), 13:1–13:19 (Dec 2015). https://doi.org/10.1145/2843948, http://doi.acm.org/10.1145/2843948
12. Joyce, J.M.: Kullback-leibler divergence. International encyclopedia of statistical science pp. 720–722 (2011)
13. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer **42**(8), 30–37 (Aug 2009). https://doi.org/10.1109/MC.2009.263, http://dx.doi.org/10.1109/MC.2009.263
14. Kumar, B.: A novel latent factor model for recommender system. JISTEM-Journal of Information Systems and Technology Management **13**(3), 497–514 (2016)
15. Lu, J., Wu, D., Mao, M., Wang, W., Zhang, G.: Recommender system application developments: a survey. Decision Support Systems **74**, 12–32 (2015)
16. Oulasvirta, A., Hukkinen, J.P., Schwartz, B.: When more is less: The paradox of choice in search engine use. In: Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 516–523. SIGIR '09, New York, NY, USA (2009). https://doi.org/10.1145/1571941.1572030, http://doi.acm.org/10.1145/1571941.1572030
17. Pazzani, M.J.: A framework for collaborative, content-based and demographic filtering. Artificial intelligence review **13**(5-6), 393–408 (1999)

18. Ross, A., Willson, V.L.: Basic and Advanced Statistical Tests: Writing Results Sections and Creating Tables and Figures. Springer (2018)
19. Su, X., Khoshgoftaar, T.M.: A survey of collaborative filtering techniques. Adv. in Artif. Intell. **2009**, 4:2–4:2 (Jan 2009). https://doi.org/10.1155/2009/421425, http://dx.doi.org/10.1155/2009/421425
20. Zhao, W.X., Li, S., He, Y., Wang, L., Wen, J.R., Li, X.: Exploring demographic information in social media for product recommendation. Knowledge and Information Systems **49**(1), 61–89 (2016)
21. Zhou, Y., Wilkinson, D., Schreiber, R., Pan, R.: Large-scale parallel collaborative filtering for the netflix prize. In: Proceedings of the 4th International Conference on Algorithmic Aspects in Information and Management. pp. 337–348. AAIM '08 (2008)